Accurate and efficient SAXS/SANS implementation including solvation layer effects suitable for molecular simulations

DIPARTIMENTO DI BIOSCIENZE

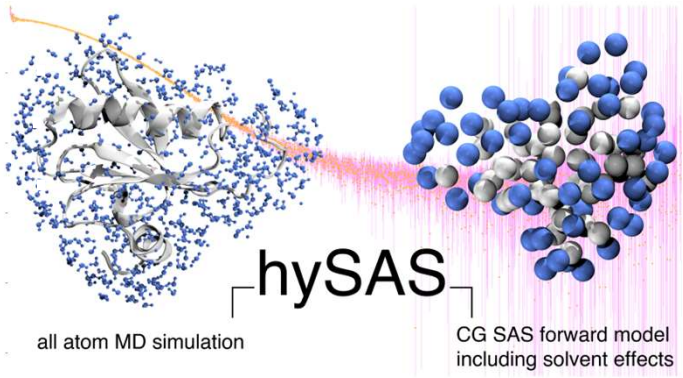
UNIVERSITÀ DEGLI STUDI DI MILANO

Federico Ballabio, Cristina Paissoni, Michela Bollati, Matteo de Rosa, Riccardo Capelli and Carlo Camilloni

federico.ballabio@unimi.it

1. Integrating experimental Small Angle Scattering (SAS) data in Molecular Dynamics (MD) simulations.

Small Angle Scattering is a low-resolution technique based on X-rays (SAXS) or neutrons (SANS) that allows the size, shape, stoichiometry, and dynamics of biomolecules to be assessed under near-physiological conditions[1].

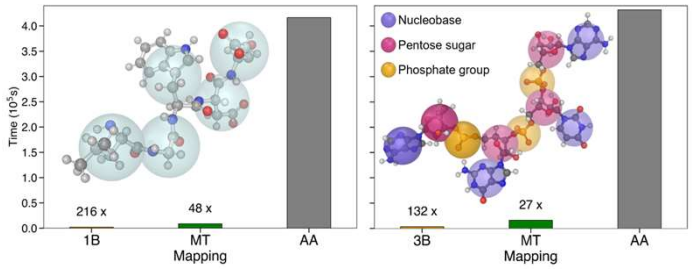


Our approach consists in restraining all-atom molecular dynamics (MD) simulations with SAS data using a coarse-grained (CG) forward model that reproduces an amino acid with a single bead (1B) and a nucleotide with three beads (3B). The beads exposed to the solvent are corrected on-the-fly to include solute-solvent scattering effects (SLC) at no additional computational cost. We call this hybrid approach hySAS and it is implemented in PLUMED[2], allowing it to be used in combination with different MD engines and restraining strategies.

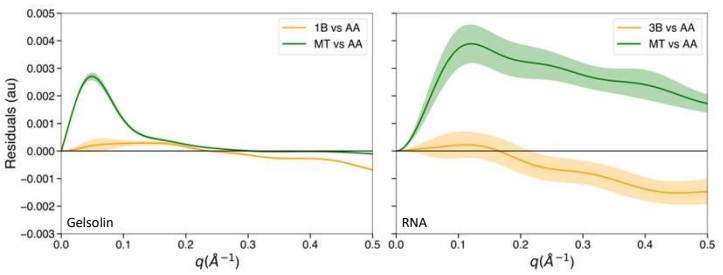
2. 1B mapping for amino acids and 3B mapping for nucleotides are fast and accurate for small *q* values.

The forward model performance is evaluated at different resolutions: all-atom (AA), Martini (MT), 1B per amino acid / 3B per nucleotide.

- 6,500 frames of gelsolin: 11,558 atoms (AA), 1,627 MT beads, 775 1B beads.
- 500 frames of ribosome RNA: 38,287 atoms, 7,796 MT beads, 3,560 3B beads



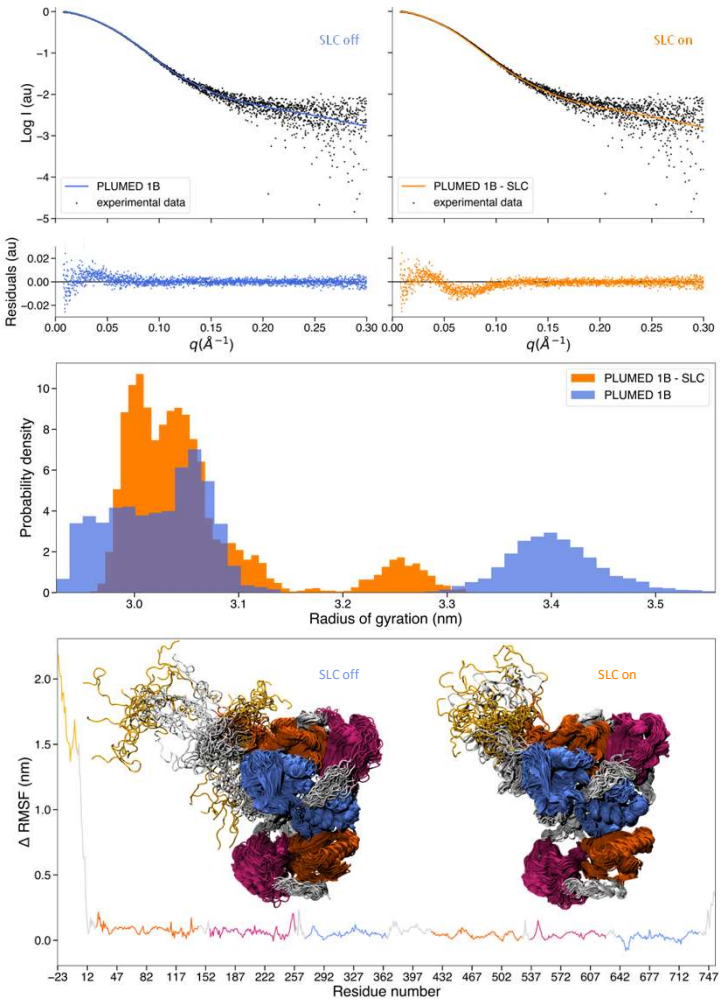
For each frame, the intensity computed with 1B/3B and with MT representation is compared with the corresponding intensity at atomistic resolution, which is taken as the reference.



The gelsolin SAXS intensities calculated with 1B mapping are in better agreement with those obtained with AA resolution than with MT up to 0.3 $Å^{-1}$. As for the proteins, the calculation of the SAXS intensity on RNA with 3B mapping also proves to be accurate, since the difference (residuals) between 3B and AA is smaller than the difference between MT and AA. These results were obtained without considering the solvation layer contribution (SLC).

3. Case study 1: determination of gelsolin conformational ensembles

Two conformational ensembles were generated via metainference multi-replica simulations, using SAXS data as restraint and 1B as forward model. One of the two ensembles was obtained by enabling the SLC. An average SAXS profile was determined from each ensemble and compared with the experimental SAXS data.

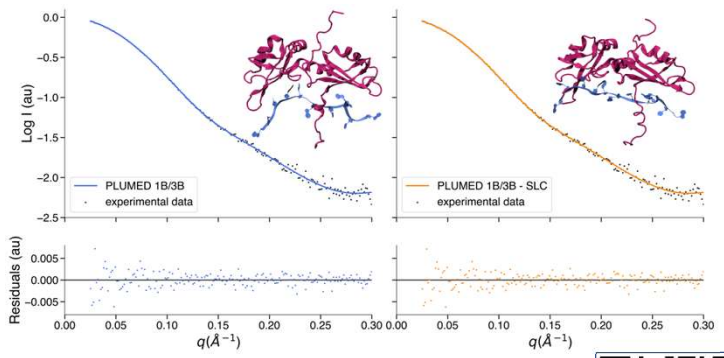


Including the SLC leads to more compact and stable models.

- SLC on: average RG of 3.05 nm, RMSF (residues) of 0.26 nm
- SLC off: average RG of 3.14 nm, RMSF (residues) of 0.38 nm

4. Case study 2: structure refinement of protein-RNA complex

The previously published structure[3] of the unwinding protein1 (UP1) interacting with a 12-mer single strand RNA has been refined to enhance consistency with experimental SAXS data.



Including the SLC leads to more compact models.

- SLC on: average RG of 2.20 nm
- SLC off: average RG of 2.26 nm

REFS: 1 Tuukkanen A.T. et al. 2017, *International Union of Crystallography Journal* | 2 Bonomi M. et al. 2016, *Science Advances* | 3 Kooshapur H. et al. 2018, *Nature Communications*